

0 W  
5635



PCT09

# ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,729B

DATE: 12/04/2002 P.6

TIME: 14:01:30

Input Set : A:\pf0687usn\_seqlist.txt

Output Set: N:\CRF4\12032002\I980729B.raw

1 <110> APPLICANT: LAL, Preeti  
 2 YUE, Henry  
 3 TANG, Y. Tom  
 4 HILLMAN, Jennifer L.  
 5 BAUGHN, Mariah R.  
 6 YANG, Junming  
 8 <120> TITLE OF INVENTION: Carbohydrate-Modifying Enzyme  
 10 <130> FILE REFERENCE: PF-0687 USN  
 12 <140> CURRENT APPLICATION NUMBER: 09/980,729B  
 C--> 13 <141> CURRENT FILING DATE: 2002-09-10  
 15 <150> PRIOR APPLICATION NUMBER: PCT/US00/10882  
 16 <151> PRIOR FILING DATE: 2000-04-20  
 18 <150> PRIOR APPLICATION NUMBER: 60/130,383  
 19 <151> PRIOR FILING DATE: 1999-04-21  
 21 <160> NUMBER OF SEQ ID NOS: 10  
 22 <170> SOFTWARE: PERL Program  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 434  
 26 <212> TYPE: PRT  
 27 <213> ORGANISM: Homo sapiens  
 29 <220> FEATURE:  
 30 <221> NAME/KEY: misc\_feature  
 31 <223> OTHER INFORMATION: Incyte ID No: 000422CD1  
 33 <400> SEQUENCE: 1  
 34 Met Asp Ser Val Glu Lys Gly Ala Ala Thr Ser Val Ser Asn Pro  
 35 1 5 10 15  
 36 Arg Gly Arg Pro Ser Arg Gly Arg Pro Pro Lys Leu Gln Arg Asn  
 37 20 25 30  
 38 Ser Arg Gly Gly Gln Gly Arg Gly Val Glu Lys Pro Pro His Leu  
 39 35 40 45  
 40 Ala Ala Leu Ile Leu Ala Arg Gly Gly Ser Lys Gly Ile Pro Leu  
 41 50 55 60  
 42 Lys Asn Ile Lys His Leu Ala Gly Val Pro Leu Ile Gly Trp Val  
 43 65 70 75  
 44 Leu Arg Ala Ala Leu Asp Ser Gly Ala Phe Gln Ser Val Trp Val  
 45 80 85 90  
 46 Ser Thr Asp His Asp Glu Ile Glu Asn Val Ala Lys Gln Phe Gly  
 47 95 100 105  
 48 Ala Gln Val His Arg Arg Ser Ser Glu Val Ser Lys Asp Ser Ser  
 49 110 115 120  
 50 Thr Ser Leu Asp Ala Ile Ile Glu Phe Leu Asn Tyr His Asn Glu  
 51 125 130 135  
 52 Val Asp Ile Val Gly Asn Ile Gln Ala Thr Ser Pro Cys Leu His

## RAW SEQUENCE LISTING

DATE: 12/04/2002

PATENT APPLICATION: US/09/980,729B

TIME: 14:01:30

Input Set : A:\pf0687usn\_seqlist.txt

Output Set: N:\CRF4\12032002\I980729B.raw

```

53          140          145          150
54 Pro Thr Asp Leu Gln Lys Val Ala Glu Met Ile Arg Glu Glu Gly
55          155          160          165
56 Tyr Asp Ser Val Phe Ser Val Val Arg Arg His Gln Phe Arg Trp
57          170          175          180
58 Ser Glu Ile Gln Lys Gly Val Arg Glu Val Thr Glu Pro Leu Asn
59          185          190          195
60 Leu Asn Pro Ala Lys Arg Pro Arg Arg Gln Asp Trp Asp Gly Glu
61          200          205          210
62 Leu Tyr Glu Asn Gly Ser Phe Tyr Phe Ala Lys Arg His Leu Ile
63          215          220          225
64 Glu Met Gly Tyr Leu Gln Gly Gly Lys Met Ala Tyr Tyr Glu Met
65          230          235          240
66 Arg Ala Glu His Ser Val Asp Ile Asp Val Asp Ile Asp Trp Pro
67          245          250          255
68 Ile Ala Glu Gln Arg Val Leu Arg Tyr Gly Tyr Phe Gly Lys Glu
69          260          265          270
70 Lys Leu Lys Glu Ile Lys Leu Leu Val Cys Asn Ile Asp Gly Cys
71          275          280          285
72 Leu Thr Asn Gly His Ile Tyr Val Ser Gly Asp Gln Lys Glu Ile
73          290          295          300
74 Ile Ser Tyr Asp Val Lys Asp Ala Ile Gly Ile Ser Leu Leu Lys
75          305          310          315
76 Lys Ser Gly Ile Glu Val Arg Leu Ile Ser Glu Arg Ala Cys Ser
77          320          325          330
78 Lys Gln Thr Leu Ser Ser Leu Lys Leu Asp Cys Lys Met Glu Val
79          335          340          345
80 Ser Val Ser Asp Lys Leu Ala Val Val Asp Glu Trp Arg Lys Glu
81          350          355          360
82 Met Gly Leu Cys Trp Lys Glu Val Ala Tyr Leu Gly Asn Glu Val
83          365          370          375
84 Ser Asp Glu Glu Cys Leu Lys Arg Val Gly Leu Ser Gly Ala Pro
85          380          385          390
86 Ala Asp Ala Cys Ser Thr Ala Gln Lys Ala Val Gly Tyr Ile Cys
87          395          400          405
88 Lys Cys Asn Gly Gly Arg Gly Ala Ile Arg Glu Phe Ala Glu His
89          410          415          420
90 Ile Cys Leu Leu Met Glu Lys Val Asn Asn Ser Cys Gln Lys
91          425          430
93 <210> SEQ ID NO: 2
94 <211> LENGTH: 302
95 <212> TYPE: PRT
96 <213> ORGANISM: Homo sapiens
97 <214> ORGANISM: Homo sapiens
98 <220> FEATURE:
99 <221> NAME/KEY: misc_feature
100 <223> OTHER INFORMATION: Incyte ID No: 983984CD1
102 <400> SEQUENCE: 2
103 Met Lys Ala Pro Gly Arg Leu Val Leu Ile Ile Leu Cys Ser Val
104 1 5 10 15

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,729B

DATE: 12/04/2002

TIME: 14:01:30

Input Set : A:\pf0687usn\_seqlist.txt

Output Set: N:\CRF4\12032002\I980729B.raw

```

105 Val Phe Ser Ala Val Tyr Ile Leu Leu Cys Cys Trp Ala Gly Leu
106                20                25                30
107 Pro Leu Cys Leu Ala Thr Cys Leu Asp His His Phe Pro Thr Gly
108                35                40                45
109 Ser Arg Pro Thr Val Pro Gly Pro Leu His Phe Ser Gly Tyr Ser
110                50                55                60
111 Ser Val Pro Asp Gly Lys Pro Leu Val Arg Glu Pro Cys Arg Ser
112                65                70                75
113 Cys Ala Val Val Ser Ser Ser Gly Gln Met Leu Gly Ser Gly Leu
114                80                85                90
115 Gly Ala Glu Ile Asp Ser Ala Glu Cys Val Phe Arg Met Asn Gln
116                95                100               105
117 Ala Pro Thr Val Gly Phe Glu Ala Asp Val Gly Gln Arg Ser Thr
118                110               115               120
119 Leu Arg Val Val Ser His Thr Ser Val Pro Leu Leu Leu Arg Asn
120                125               130               135
121 Tyr Ser His Tyr Phe Gln Lys Ala Arg Asp Thr Leu Tyr Met Val
122                140               145               150
123 Trp Gly Gln Gly Arg His Met Asp Arg Val Leu Gly Gly Arg Thr
124                155               160               165
125 Tyr Arg Thr Leu Leu Gln Leu Thr Arg Met Tyr Pro Gly Leu Gln
126                170               175               180
127 Val Tyr Thr Phe Thr Glu Arg Met Met Ala Tyr Cys Asp Gln Ile
128                185               190               195
129 Phe Gln Asp Glu Thr Gly Lys Asn Arg Arg Gln Ser Gly Ser Phe
130                200               205               210
131 Leu Ser Thr Gly Trp Phe Thr Met Ile Leu Ala Leu Glu Leu Cys
132                215               220               225
133 Glu Glu Ile Val Val Tyr Gly Met Val Ser Asp Ser Tyr Cys Arg
134                230               235               240
135 Glu Lys Ser His Pro Ser Val Pro Tyr His Tyr Phe Glu Lys Gly
136                245               250               255
137 Arg Leu Asp Glu Cys Gln Met Tyr Leu Ala His Glu Gln Ala Pro
138                260               265               270
139 Arg Ser Ala His Arg Phe Ile Thr Glu Lys Ala Val Phe Ser Arg
140                275               280               285
141 Trp Ala Lys Lys Arg Pro Ile Val Phe Ala His Pro Ser Trp Arg
142                290               295               300
143 Thr Glu
146 <210> SEQ ID NO: 3
147 <211> LENGTH: 578
148 <212> TYPE: PRT
149 <213> ORGANISM: Homo sapiens
151 <220> FEATURE:
152 <221> NAME/KEY: misc_feature
153 <223> OTHER INFORMATION: Incyte ID No: 2210054CD1
155 <400> SEQUENCE: 3
156 Met Pro Phe Arg Leu Leu Ile Pro Leu Gly Leu Leu Cys Ala Leu
157 1                5                10                15

```

## RAW SEQUENCE LISTING

DATE: 12/04/2002

PATENT APPLICATION: US/09/980,729B

TIME: 14:01:30

Input Set : A:\pf0687usn\_seqlist.txt

Output Set: N:\CRF4\12032002\I980729B.raw

158	Leu	Pro	Gln	His	His	Gly	Ala	Pro	Gly	Pro	Asp	Gly	Ser	Ala	Pro
159					20					25					30
160	Asp	Pro	Ala	His	Tyr	Arg	Glu	Arg	Val	Lys	Ala	Met	Phe	Tyr	His
161					35					40					45
162	Ala	Tyr	Asp	Ser	Tyr	Leu	Glu	Asn	Ala	Phe	Pro	Phe	Asp	Glu	Leu
163					50					55					60
164	Arg	Pro	Leu	Thr	Cys	Asp	Gly	His	Asp	Thr	Trp	Gly	Ser	Phe	Ser
165					65					70					75
166	Leu	Thr	Leu	Ile	Asp	Ala	Leu	Asp	Thr	Leu	Leu	Ile	Leu	Gly	Asn
167					80					85					90
168	Val	Ser	Glu	Phe	Gln	Arg	Val	Val	Glu	Val	Leu	Gln	Asp	Ser	Val
169					95					100					105
170	Asp	Phe	Asp	Ile	Asp	Val	Asn	Ala	Ser	Val	Phe	Glu	Thr	Asn	Ile
171					110					115					120
172	Arg	Val	Val	Gly	Gly	Leu	Leu	Ser	Ala	His	Leu	Leu	Ser	Lys	Lys
173					125					130					135
174	Ala	Gly	Val	Glu	Val	Glu	Ala	Gly	Trp	Pro	Cys	Ser	Gly	Pro	Leu
175					140					145					150
176	Leu	Arg	Met	Ala	Glu	Glu	Ala	Ala	Arg	Lys	Leu	Leu	Pro	Ala	Phe
177					155					160					165
178	Gln	Thr	Pro	Thr	Gly	Met	Pro	Tyr	Gly	Thr	Val	Asn	Leu	Leu	His
179					170					175					180
180	Gly	Val	Asn	Pro	Gly	Glu	Thr	Pro	Val	Thr	Cys	Thr	Ala	Gly	Ile
181					185					190					195
182	Gly	Thr	Phe	Ile	Val	Glu	Phe	Ala	Thr	Leu	Ser	Ser	Leu	Thr	Gly
183					200					205					210
184	Asp	Pro	Val	Phe	Glu	Asp	Val	Ala	Arg	Val	Ala	Leu	Met	Arg	Leu
185					215					220					225
186	Trp	Glu	Ser	Arg	Ser	Asp	Ile	Gly	Leu	Val	Gly	Asn	His	Ile	Asp
187					230					235					240
188	Val	Leu	Thr	Gly	Lys	Trp	Val	Ala	Gln	Asp	Ala	Gly	Ile	Gly	Ala
189					245					250					255
190	Gly	Val	Asp	Ser	Tyr	Phe	Glu	Tyr	Leu	Val	Lys	Gly	Ala	Ile	Leu
191					260					265					270
192	Leu	Gln	Asp	Lys	Lys	Leu	Met	Ala	Met	Phe	Leu	Glu	Tyr	Asn	Lys
193					275					280					285
194	Ala	Ile	Arg	Asn	Tyr	Thr	Arg	Phe	Asp	Asp	Trp	Tyr	Leu	Trp	Val
195					290					295					300
196	Gln	Met	Tyr	Lys	Gly	Thr	Val	Ser	Met	Pro	Val	Phe	Gln	Ser	Leu
197					305					310					315
198	Glu	Ala	Tyr	Trp	Pro	Gly	Leu	Gln	Ser	Leu	Ile	Gly	Asp	Ile	Asp
199					320					325					330
200	Asn	Ala	Met	Arg	Thr	Phe	Leu	Asn	Tyr	Tyr	Thr	Val	Trp	Lys	Gln
201					335					340					345
202	Phe	Gly	Gly	Leu	Pro	Glu	Phe	Tyr	Asn	Ile	Pro	Gln	Gly	Tyr	Thr
203					350					355					360
204	Val	Glu	Lys	Arg	Glu	Gly	Tyr	Pro	Leu	Arg	Pro	Glu	Leu	Ile	Glu
205					365					370					375
206	Ser	Ala	Met	Tyr	Leu	Tyr	Arg	Ala	Thr	Gly	Asp	Pro	Thr	Leu	Leu

## RAW SEQUENCE LISTING

DATE: 12/04/2002

PATENT APPLICATION: US/09/980,729B

TIME: 14:01:30

Input Set : A:\pf0687usn\_seqlist.txt

Output Set: N:\CRF4\12032002\I980729B.raw

```

207          380          385          390
208 Glu Leu Gly Arg Asp Ala Val Glu Ser Ile Glu Lys Ile Ser Lys
209          395          400          405
210 Val Glu Cys Gly Phe Ala Thr Ile Lys Asp Leu Arg Asp His Lys
211          410          415          420
212 Leu Asp Asn Arg Met Glu Ser Phe Phe Leu Ala Glu Thr Val Lys
213          425          430          435
214 Tyr Leu Tyr Leu Leu Phe Asp Pro Thr Asn Phe Ile His Asn Asn
215          440          445          450
216 Gly Ser Thr Phe Asp Thr Val Ile Thr Pro Tyr Gly Glu Cys Ile
217          455          460          465
218 Leu Gly Ala Gly Gly Tyr Ile Phe Asn Thr Glu Ala His Pro Ile
219          470          475          480
220 Asp Pro Ala Ala Leu His Cys Cys Gln Arg Leu Lys Glu Glu Gln
221          485          490          495
222 Trp Glu Val Glu Asp Leu Met Arg Glu Phe Tyr Ser Leu Lys Arg
223          500          505          510
224 Ser Arg Ser Lys Phe Gln Lys Asn Thr Val Ser Ser Gly Pro Trp
225          515          520          525
226 Glu Pro Pro Ala Arg Pro Gly Thr Leu Phe Ser Pro Glu Asn His
227          530          535          540
228 Asp Gln Ala Arg Glu Arg Lys Pro Ala Lys Gln Lys Val Pro Leu
229          545          550          555
230 Leu Ser Cys Pro Ser Gln Pro Phe Thr Ser Lys Leu Ala Leu Leu
231          560          565          570
232 Gly Gln Val Phe Leu Asp Ser Ser
233          575
235 <210> SEQ ID NO: 4
236 <211> LENGTH: 461
237 <212> TYPE: PRT
238 <213> ORGANISM: Homo sapiens
240 <220> FEATURE:
241 <221> NAME/KEY: misc_feature
242 <223> OTHER INFORMATION: Incyte ID No: 2618358CD1
244 <400> SEQUENCE: 4
245 Met Gly Gly Ser Thr Ala Ala Tyr Gln Val Glu Gly Gly Trp Asp
246 1          5          10          15
247 Ala Asp Gly Lys Gly Pro Cys Val Trp Asp Thr Phe Thr His Gln
248          20          25          30
249 Gly Gly Glu Arg Val Phe Lys Asn Gln Thr Gly Asp Val Ala Cys
250          35          40          45
251 Gly Ser Tyr Thr Leu Trp Glu Glu Asp Leu Lys Cys Ile Lys Gln
252          50          55          60
253 Leu Gly Leu Thr His Tyr Arg Phe Ser Leu Ser Trp Ser Arg Leu
254          65          70          75
255 Leu Pro Asp Gly Thr Thr Gly Phe Ile Asn Gln Lys Gly Ile Asp
256          80          85          90
257 Tyr Tyr Asn Lys Ile Ile Asp Asp Leu Leu Lys Asn Gly Val Thr
258          95          100          105

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/980,729B

DATE: 12/04/2002  
TIME: 14:01:31

Input Set : A:\pf0687usn\_seqlist.txt  
Output Set: N:\CRF4\12032002\I980729B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 1398

**VERIFICATION SUMMARY**

DATE: 12/04/2002

PATENT APPLICATION: US/09/980,729B

TIME: 14:01:31

Input Set : A:\pf0687usn\_seqlist.txt

Output Set: N:\CRF4\12032002\I980729B.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:470 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:1380